Asian Journal of Multidimensional Research
ISSN: 2278-4853 Vol. 11, Issue 10, October 2022 SJIF 2022 = 8.179
A peer reviewed journal

A NEURO-GENETIC FUZZY SYSTEM FOR THE PREDICTION OF STUDENT'S ACADEMIC PERFORMANCE

Ojekudo*; Nathaniel Akpofure**

*(Ph.D),
Department of Computer Science,
Ignatius Ajuru University of Education,
Nigeria,
Email id: nathojekudo@gmail.com

DOI: 10.5958/2278-4853.2022.00269.5

ABSTRACT

The Nigerian education sector has over the years being facing serious problems of producing half-baked graduates or what others termed to be "educated illiterates" in various fields of human endeavour. The root cause of this issue is traced to poor academic performance evaluation of students seeking admission into tertiary institutions. The current model for evaluating the academic performance of admission seekers encompasses the Joint Admission and Matriculation Board (JAMB) and the West African Senior School Certificate Examinations (WASSCE), National Examination Council (NECO) or even National Board of Technical Education Council (NBTE). However, it has being argued that rather than JAMB and WASSCE or its equivalent, to evaluate admission seekers based on excellent correlated academic performances, the model has now become a citadel of bribery, corruption and gross academic malpractices. A model was developed to address this problem using tree based ensemble machine learning algorithms but we discovered that it did not sufficiently address the problem. In this work, we developed an Enhanced Neuro-Genetic Fuzzy Model for predicting students' academic performance using structured analysis and design methodology. The Neuro-Genetic was used for training and optimizing the proposed model. Furthermore, we utilized 1000 datasets of which 60%, @0% and 20% was used for training testing and validation of the proposed model. We implemented the proposed system with Java programming language and MySQL relational database as backend. Our obtained result showed better Grand Mean (x) of users' satisfaction when compared to the existing system result in terms of graphical user interface, accuracy of the prediction process, speed of the prediction process, speed in user validation and security. The proposed system Grand Mean obtained after comparative analysis are 2.74%, 1.78%, 2.71%, 1.67%, and 2.57% respectively, while those obtained by the existing system are 2.59%, 0.78%, 2.68%, 0.67%, and 2.52% respectively. In addition, the study also recommended the need for more application of machine-learning concepts to the prediction of student's academic performance. Machine learning techniques such as neuro-genetic fuzzy have proven to be efficient in predicting the academic performance of students.

KEYWORDS: *Academic Performance, Fuzzy, Genetic, Neuro, Prediction*

I. INTRODUCTION

The study addresses the impact of half-baked graduates which further results in corrosive academic mindsets and myopic intelligence application among majority of churned-out graduates in Nigeria. The root cause of this issue is traced to poor academic performance evaluation of students seeking admission into tertiary institutions. The current model for evaluating the academic performance of admission seekers encompasses the Joint Admission and Matriculation Board (JAMB) and the West African Senior School Certificate Examinations (WASSCE), National Examination Council (NECO) or even National Board of Technical Education Council (NBTE). However, it has being argued that rather than JAMB and WASSCE or its equivalent, to evaluate admission seekers based on excellent correlated academic performances, the model which has now become a citadel of bribery, corruption and gross academic malpractices, there should be a better model.

According to Yusuf (2019), "The Nigerian education sector has over the years being facing serious problems of producing half-baked graduates or what others termed to be "educated illiterates" in various fields of human endeavor". This in essence translates that tertiary institutions will continue to produce graduates with low level of intelligence to handle the affairs of our respective societies and of course in different endeavor including leadership and nation-building except something is done drastically. Student academics performance has been a problem to institutional manager, government and parents. In other to make a choice of capable and qualified students to admit, a lot of examinations have to be taken. Examination bodies have been tried to filter the best qualified students to get admitted. Functions have been duplicated by these examination bodies, all in bids to getting best qualified students who are expected to be exceptional and become good professionals that the society may need tomorrow. This have not only yielded very disappointing results but left the stakeholders in a more confusing state. Performance may be defined as a measurable behavior in a particular situation. Despite the number of entry examination, students take in other to gain admission in universities, the rate of "drop out" or failing students are much. It is even very glaring that the number of "Alpha" in their WAEC and NECO examinations coupled with their high JAMB score do not necessarily make them become brilliant students. Making a choice of the best way of selecting new entrance into the university has caused so much controversy and crises among stakeholders of the educational sector. This is as a result of poor performance by students of our higher institution. Blames has been placed on certain levels as laying poor academic foundations for other. Some stakeholders blame the processes of selection into these institutions, believing that assessment was meant to be done to choose only the very best. The question now is "How successful is the assessment?"

An early prediction of students' academic performance is what this assessment tends to pursue. This is supposed to be a very important practice as it tries to obtain an ideal students' level of learning, level of teaching, decides the success or failure of students in the course of enrolment, to inform low preforming students to put in more effort and overcome weaknesses and encourages students to either continue or choose an appropriate course of carrier. Different institutions of learning have diverse ways of admitting new students. Some consider cognitive ability and personality while others admit based on merit. All these may not yield desired results. A multi-fictile approach that considers more constraints and variables, taking cognizance of data may yield a better result.

The rate of final cumulative of such students first becomes necessary to study the obvious very important parameter for entry into these institutions and see whether the right parameters are being used. This may address the global call for reduction of mass failure and drop out in our institutions of higher learning. It may also address the "pass by all means" syndrome and the high rate of examination malpractices currently seen in our society. The prediction of students' academic performance had been carried out using different approaches with consequently different kinds of results but into a common purpose. Models such Fuzzy Probabilistic Neural Network, Heuristic Models, Logistic Regression Ensemble, Multi-Gene Genetic Programming Application, Fuzzy Logic, Neuro-Fuzzy, Rules Optimization Based Fuzzy Models and so many more have been used for one prediction and the other.

This study adopts a Hybridized Model for academic performance prediction which encompasses Deep Learning, Fuzzy Logic and Genetic Algorithm (i.e. Deep Fuzzy-Genetic). Deep learning can be described as a technique which uses neural networks for the training of data in order to build a model. Fuzzy logic is a branch of machine learning that is being utilized to determine partial truth of a computing process using Boolean value which includes 0 and 1. Fuzzy logic algorithm helps to solve a problem after considering all available data. Then it takes the best possible decision for the given input. It is a concept that is still being applied in industrial areas such as manufacturing, electricity, robotics, etc.

Genetic Algorithm (GA) is a selection algorithm used for solving complex search optimization issues. Genetic Algorithm is tailored after the principle of sexual reproduction and survival of chromosomes as it concerns sex of a child. It is simply a search algorithm which is based on survival of the fittest in competing for resources in an environment. The idea tends to maintain population of chromosomes, as in human, but in this case, represents candidate solutions to a problem and the candidate will evolve over a period of time through competition and controlled variation.

1.1 Statement of the Problem

The study intends to address several factors that negatively influence the academic performance of students in tertiary institutions. Factors that negatively affect students' academic performance include poor motivation from lecturers, bad peer group, poor teaching quality, etc. Furthermore, the mandatory requirement for admission into the first year degree programmes for all Nigerian Universities and Tertiary institutions are the combination of JAMB and WASSCE. The integrity of the mentioned examination has now become doubtful. Hence, most universities had introduced the Post UTME (Unified Tertiary Matriculation Examination) examinations to further assess their prospective students. However, the integrity of Post-UTMEs has also been subjected to public scrutiny especially in cases which involve corruption and gross examination malpractices. The mentioned problems have continued to exist because there is no enhanced machine-learning model for predicting the academic performances of students that are seeking admission into tertiary institutions.

Henry et al (2019) came close to addressing the mentioned problems through the prediction of school performance using a combination of traditional and non-traditional data from South Africa. The study presented a preliminary research and describes the importance of tree-based machine learning algorithms and their applications. They adopted the combination of models such as the LightGBM, XGBoost, Random Forest and Decision Trees. In order to effectively

optimize the aim of their work, the authors explored different nontraditional data sources from previous examination results, school features such as location, water availability, internet, household goods and municipality difficulties in South Africa.

Furthermore, the authors built tree based interpretable models that apply 'if-then' rules as they are easy to be interpreted by policy makers. However, the adopted tree-based models had limitations. They require correct parameter tuning. The LightGBM leaf-wise growth may be over-fitting if wrong parameters are used. For example, large number of leaves may cause over-fitting and parameters are dependent on each other. Also, the adopted models in the study did not consider individual learning factors such as motivation, intellectual ability or prior knowledge. In other words, this study intends to address two major problems which include latencies in accurate prediction of students' academic performance due to over-fitting of the trees-based models adopted by Henry et al (2019), and also the absence of individual learning factors such as motivation, intellectual ability or prior knowledge which are supposed to be used as build-up tools for accurate prediction of students' academic performances.

1.2 Aim and Objectives of the Study

The aim of this study is to develop a Deep Neuro-Genetic Fuzzy System for the Prediction of Students Academic Performance. The specific objectives of the study include to:

i) Design an Academic Model for predicting Students' Performance

ii) Train the proposed Academic Model with Deep Neural Network

iii) Optimize the proposed Academic Model with Fuzzy-Genetic Technique

iv) Implement the proposed Academic Model with Java Programming Language

v) Evaluate performance of the proposed Academic Model using appropriate evaluation metrics

1.3 Significance of the Study

The benefit of this study encompasses a structured correlated framework for accurately predicting the academic performances for seekers of admission into tertiary institutions. In other words, the study will be beneficial to the following:

i) The Management of Tertiary Institutions in Nigeria

ii) Seekers of Admission into Tertiary Institutions.

iii) Researchers with keen interest in Education Data Mining (EDM)

II. RELATED WORKS

Henry et al, (2019), looked at predicting school performance using a combination of traditional and non-traditional data from South Africa. The study presented a preliminary research and describes the importance of tree-based machine learning algorithms and their application. They adopted the combination of models such as the Light GBM, XG Boost, Random Forest and Decision Trees. In order to effectively optimize the aim of their work, the authors explored different nontraditional data sources from previous examination results, school features such as location, water availability, internet, household goods and municipality difficulties in South

Africa. Furthermore, the authors built tree based interpretable models that apply 'if-then' rules as they are easy to be interpreted by policy makers. The LightGBM model underlies the developed final prediction model as it outperformed traditional tree based models such as XGboost and random forest. In addition, the authors used SHAP (SHapley Additive exPlanations) to explain model outputs and provide relationships among variables to get better understanding of education systems and the underlying factors influencing performance in schools. The authors further compared the performance of LightGBM with the prediction performance of the other benchmark models trained using XGBoost, random forest and decision trees. Results from the study show that the LightGBM model emerged best with its results being comparable to XGBoost. The authors did a good job. However, their adopted tree-based models had limitations. They require correct parameter tuning. The LightGBM leaf-wise growth may be over-fitting if wrong parameters are used, for example, large number of leaves may cause over-fitting and parameters are dependent on each other. Also, the adopted models in the study did not consider individual learning factors such as motivation, intellectual ability or prior knowledge.

Kuyoro et al, (2013), proposed an optimal algorithm for predicting the academic performance which includes the senior certificate examination, unified matriculation examination and first year's cumulative grade point respectively. The model recorded some degree of accuracy but was impeded by the dataset employed which may not portray the true picture of the students' baseline performance. And on this premise the aim of the study may not be fully actualized as student with high Entry Grade point may not be amongst the best performing in the group thereby increasing the false portives. It is therefore debatable that entry requirements and first year grades point may not be the best guide for measuring performance irrespective of family background.

Arora and Saini (2014), proposed a Neuro-Fuzzy system which modeled students' performance mimicking the teachers' ability to linguistically manage uncertainty in the bid to increase the degree of correctness and accuracy of the model using MATLAB for the simulation process. The drawback of this process irrespective of the degree of accuracy recorded the inefficiency of the model was tied to the errors in the modeling process as the teacher's approach to predicting performance was defective ab initio, because if it was taken into consideration the accuracy of the model would have surpassed the results obtained.

Amrieh et al, (2016), carried out a comparative review of different data mining algorithms for predicting performance of students based on behavioral pattern using artificial Neural Networks, Naïve Bayesian and Decision Tree and ensemble methods like Bagging Booster and Random Forest to enhance the performance of the classifiers. The level of uncertainty involved with monitoring and modeling human behavior was not put into consideration as the accuracy obtained by the model would have greatly improved if a fuzzy system was deployed.

Badra et al, (2016), implemented the classification based on association (CBA) for predicting students' performance with the aim of eliminating the high rate of drop outs by proffering remedial measures to enhance students' performance beforehand. The model performance was low compared to other data mining algorithms as it was unable to address some critical issues that borders on the initial or entry qualification.

Imran et al, (2019), proposed a supervised learning decision tree classifier for predicting students' performance using ensembling methods to enhance the performance of the classifier.

The system had need to be fine-tuned and continuous data preprocessing to achieve optimal result, secondly the problem of predicting performance could in some cases not be fully optimized using supervised methods as it tends produce a stereotyped result that may not be practicable in real life.

Gbolagade et al, (2015), deployed an artificial neural network model with Rensample Preprocessed Algorithm in WEKA for the prediction of the performance of postgraduate students in University of Ilorin. Using radial basis function and multilayer perceptron the model predicted the performance of students for eight sessions taking into consideration demographic factors and which in turn caused the model to over fit with low performance. The parameters used for implementing the model was in most cases unnecessary as it affected its performance.

Chandrashekhar et al., (2015), researched on clustering-based feature selection algorithm. They combined active feature selection, genetic algorithm and bisecting K-means to describe the hybrid method used for text clustering. They also used business intelligence to get some guided data mining methods by identifying the related services. They compared their hybrid method with K-means and their results showed that their hybrid method performed better than K-means in terms of accuracy and clustering results. But they could not explore other data mining algorithms such association and classification algorithms.

Ismaila et al., (2015), presented a combined a negative selection algorithm-particle swarm optimization for an academic performance detection system. In their work, they proposed a model called NSA-PSO (negative selection algorithm-particle swarm optimization) model. This model took advantage of the strengths of both the negative selection algorithm and particle swarm optimization to achieve optimal results. When they compared their NSA-PSO model with the standard NSA model, their results showed that their model had a better accuracy than the standard NSA model. But their model could not implement a side by side hybridization of two evolutionary algorithms to perform the singular task of detecting poor academic performance in their model network.

Okesola et al., (2015), used locality to improve SVM (Support Vector Machine) - based education data filtering. Their work was centered on using SVM to solve the problem of advanced free fraud. SVM was used to build local decision rules into the classification process of the data filter design for use for the academic performance of students. The result of their work showed a high level of classification accuracy. But they could not use other data filtering techniques like roulette wheel of the genetic algorithm.

Sao and Prashanthi (2015), proposed academic performance classification using naïve Bayesian classifier. The dataset they used is EduPlus Records and they used feature extraction technique for extracting the features from the students' records. In their experiment, they compared the performance of naïve Bayes with support vector machine in terms of error rate and accuracy. The result of their experiment shows that naïve Bayes is more accurate and has a lesser error rate than support vector machine. But they could not perform feature selection.

III. ANALYSIS OF THE EXISTING SYSTEM

The Existing System to improve in the study is a Hybridized Model of LightGBM, XGBoost, Random Forest and Decision Trees for predicting students' academic performance (Figure 3.1). The Hybridized Model was developed by Henry et al (2019) and described the importance of

tree-based machine learning algorithms and their application. In order to effectively optimize the aim of their work, the authors explored different nontraditional data sources from previous examination results, school features such as location, water availability, internet, household goods and municipality difficulties in South Africa. Furthermore, the authors built tree based interpretable models that apply 'if-then' rules as they are easy to be interpreted by policy makers.

The Light GBM model underlies the developed final prediction model as it outperformed traditional tree based models such as XGboost and random forest. In addition, the authors used SHAP (SHapley Additive exPlanations) to explain model outputs and provide relationships among variables to get better understanding of education systems and the underlying factors influencing performance in schools.

3.1.1 Explanation of the Existing System Components

The following components of the Existing System are:

i) Datasets

This component illustrates several data contents that were utilized for developing the existing system which encompasses school performance, location, survey and so on.

ii) Models

This component illustrates several expert system techniques behind the performance of the existing system. Furthermore, this component is made up of other sub-components such as LightGBM, Xgboost, Decision Trees, and Random Forest. The LightGBM underlies the developed final prediction model as it outperformed traditional tree based models such as XGboost and random forest.

iii) SHAP

This component illustrates SHapley Additive explanation which is used to explain model outputs and provide relationships among variables to get better understanding of education systems and the underlying factors influencing performance in schools.

iv) Prediction and Comparison

This component illustrates the testing, forecasting and selection of the best model for extracting specific academic performance of the students.

v) Interpretation

This component gives the final selection result of the best model suitable for forecasting the academic performance of students.

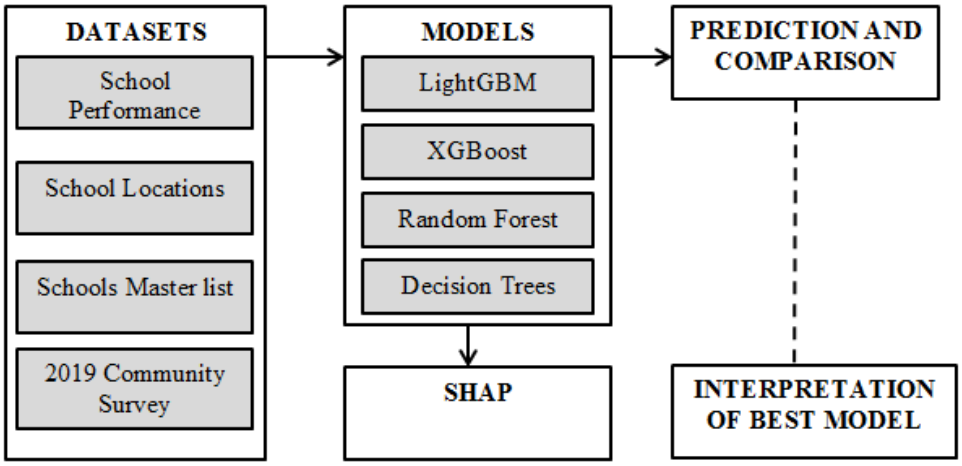

Figure 3.1: Existing System Architecture for Predicting Students' Academic Performance: (Source: Henry et al, 2019)

3.1.2 Algorithm of the Existing System

The following is the existing system algorithm

Step 1: Start
Step 2: Initialize System
Step 3: Input set of classes in program being refactored
Step 4: Input set of 4 datasets types (e.g. null-up method)
Step 5: Input set of 4 models to analyze datasets
Step 6: Output datasets and models accepted and initialized
Step 7: Initialize SHAP
Step 8: refactoring_count = 0
Step 9: repeat
Step 10: classes = set of classes in program
Step 11: while !empty (classes) do
Step 12: class = classes.pick()
Step 13: if fitness_function_improves () then
Step 14: refactoring_count++
Step 15: update system output
Step 16: else
Step 17: refactoring.undo()

Step 18: end

3.1.3 Disadvantages of the Existing System

The existing system model has some drawbacks that can affect its performance on academic data prediction. The following drawbacks are:

i) The adopted tree-based models of the existing system had limitations. They require correct parameter tuning.

ii) The LightGBM leaf-wise growth may be over-fitting if wrong parameters are used, for example, large number of leaves may cause over-fitting and parameters are dependent on each other.

iii) The adopted models in the study did not consider individual learning factors such as motivation, intellectual ability or prior knowledge.

IV. DESIGN OF THE NEW SYSTEM

The new system is an improved version of the existing system. The New System is a Deep Neuro-Genetic Fuzzy Model for the Prediction of Students' Academic Performance (Figure 3.2). The New system considered individual learning factors such as motivation, intellectual ability or prior knowledge. Furthermore, it is a hybridized model for academic performance prediction which encompasses Deep Learning, Fuzzy Logic and Genetic Algorithm (i.e. Deep Fuzzy-Genetic). Deep learning is a subset of machine learning in artificial intelligence (AI) that has networks capable of learning unsupervised data that is unstructured or unlabeled. Deep learning has evolved hand-in-hand with the digital era, which has brought about an explosion of data in all forms and from every region of the world. This data, known simply as big data, is drawn from sources like social media, internet search engines, e-commerce platforms, and online cinemas, among others. This enormous amount of data is readily accessible and can be shared through fintech applications like cloud computing.

Fuzzy logic is a branch of machine learning that is being utilized to determine partial truth of a computing process using Boolean value which includes 0 and 1. Fuzzy logic algorithm helps to solve a problem after considering all available data. Then it takes the best possible decision for the given the input. It is a concept that is still being applied in industrial areas such as manufacturing, electricity, robotics, etc. Genetic Algorithm is a machine learning algorithm that imitates the process of natural selection. It is used for solving complex search optimization issues. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation.

3.2.1 Tools for Building the New System

The study intends to improve the adopted framework of Henry et al., (2019), through the addition of a deep neuro-genetic fuzzy system for the prediction of students' academic performance in their final year. The New system is mainly a hybrid of deep learning using neural networks for an unsupervised approach to training datasets, genetic algorithm and fuzzy logic. Furthermore, the fusion of Genetic Algorithms and Fuzzy Logic Systems allows researchers to model real world problems through the development of intelligent and adaptive systems.

3.2.2 Explanation of the New System Components

This process is shows major tools such as the Genetic Algorithm Learning process, the knowledge-base, the input interface, the fuzzy rule-based system and the output interface. This process aids learning of the inputted datasets. The genetic algorithm differs from a classical, derivative-based, optimization algorithm in two essential ways. Namely, a classical algorithm generates single point iteration and the sequence of points approaches an optimal solution. The Fuzzy Logic in the design process is a system which consists of a knowledge base, which includes the information given in the form of linguistic control rules, and a fuzzification interface, which has the effect of transforming crisp data into fuzzy sets. Additionally, an Inference System is included which works together with the knowledge base to provide inference with the use of a reasoning method, and a defuzzification interface, which translates the fuzzy control action obtained to a real control action by means of a defuzzification method.

3.2.3 Equations of the New System

To ensure integrity and maintain unique instances to each prediction of student's academic performance, a MATLAB script was written to extract unique instances for the prediction process. Datasets were collected from academic records in order to formulate the model.

The Fuzzy Model

An Interval Type-2 Fuzzy Set (IT2FS) characterized by $\tilde{A}$ has a FOU bounded by a lower and upper membership functions, $\underline{\mu}_{\tilde{A}}(x,\mu)$ and $\underline{\mu}_{\tilde{A}}(x,\mu)\forall x \in X$ respectively, is expressed as:

$$\tilde{A} = \left\{\left((x,\mu), \underline{\mu}_{\tilde{A}}(x,\mu), \bar{\mu}_{\tilde{A}}(x,\mu)\right) \middle| \forall x \in X, \forall \mu \in J_x \subseteq [0,1]\right\}, \quad (3.1)$$

Where, $\underline{\mu}_{\tilde{A}}(x,\mu)$ and $\bar{\mu}_{\tilde{A}}(x,\mu)\forall x \in X$ =1

$x \in X$ and $\mu \in J_x$ are defined as a continuous universe of discourse (UoD);

x denotes the primary variable in domain X

μ denotes the secondary variable in domain J_x at each $x \in X$

J_x is called the primary membership of x as defined in (1), which symbolizes the interval set.

The secondary grades of $\tilde{A}$ is unity, thus reduces IT2FS to:

$$\tilde{A} = \int_{x\in X} \int_{\mu\in J_x} {}^{1}/_{(x,\mu)}. \quad (3.2)$$

Now, the FOU of $\tilde{A}$ is the union of all primary membership grades and is given by:

$\text{FOU}(\tilde{A}) = \bigcup_{x\in X} J_x$,

The Upper Membership Function (UMF) $\bar{\mu}_{\bar{A}}(x)$ and Lower Membership Function (LMF) $\underline{\mu}_{\bar{A}}(x)$ are type -1 membership functions (MFs) marking the FOU boundary of interval type-2 MF. The UMF represents the subset that has the maximum membership grade of the FOU; and the LMF is a subset that has the minimum membership grade of the FOU $\forall x \in X$.

Thus

$$\bar{\mu}_{\bar{A}}(x) \equiv \overline{FOU(\tilde{A})}, \forall x \in X,$$
$$\underline{\mu}_{\bar{A}}(x) \equiv \underline{FOU(\tilde{A})}, \forall x \in X, \quad (3.3)$$
$$J_x = \left[\bar{\mu}_{\bar{A}}(x)\,,\ \underline{\mu}_{\bar{A}}(x)\right].$$

The triangular membership function (TMF) was adopted to evaluate each input and output MFs for the IT2FL system. The description of the TMF using a line or curve is based on three parameters a_1, p and a_2 , and specifies the mapping of each input or output parameters, to obtain membership values for n membership grades MG_n; n: 1, …, n

Thus,

$$\mu(x) = \begin{cases} 0; if\ x < a_1(MG_1)\{NIR\} \\ \frac{x - a_1(MG_1)}{a_2(MG_1) - a_1(MG_1)}; if\ a_1(MG_1) \le x < a_2(MG_1) \\ \frac{a_2(MG_1)\ \ -x}{a_2(MG_1) - a_1(MG_2)}; if\ a_1(MG_2) \le x < a_2(MG_2), \\ \ldots \\ \frac{a_2(MG_n)\ \ -x}{a_2(MG_n) - a_1(MG_n)}; if\ a_1(MG_n) \le x < a_2(MG_n) \\ 0; if\ x\ \ge\ a_2\{NIR\} \end{cases}$$
(3.4)

Where a_1 and a_2 are the triangular end points defined by the FOU – region consisting of all the points of primary membership of elements, and NIR signifies values that are not in range. The equations illustrate a triangular shape IT2FLS with its principal T1FS, showing the end point, and P, the triangular peak location.

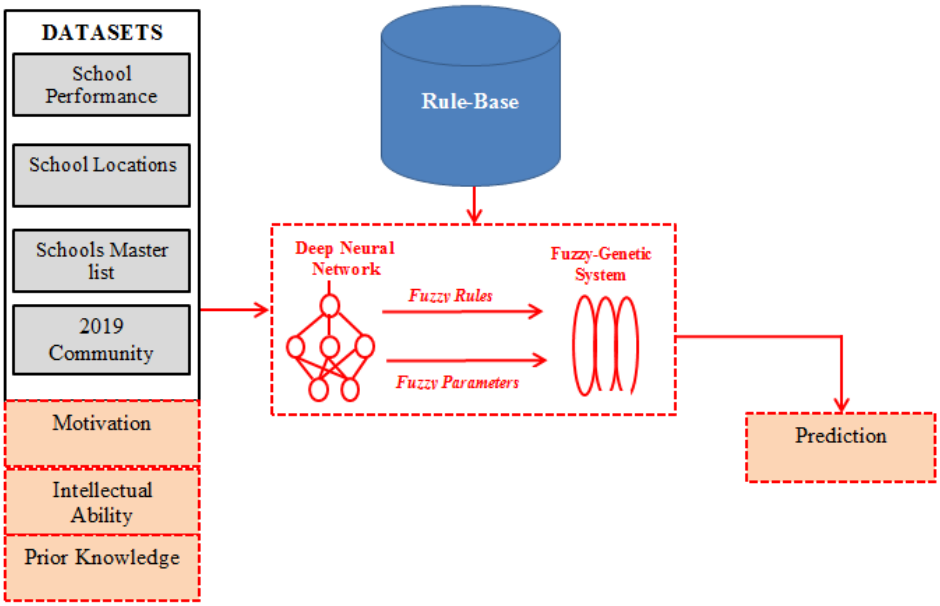

Figure 3.2: Architectural Design of the New System

3.2.4 Hybridized Algorithm of the New System

The following is the new system algorithm

Step 1: Start

Step 2: Extract Processed Deep Fuzzy-Genetic Features

Input: RB is the set of prediction rules

D is the set of documents

Output: F is the set of slot fillers extracted

Function: Information Extraction (RB,D)

F = 0

For each example DϵD do

For each rule R, RB do

If R fires on the extracted fillers

Extract the predicted filler and add it to F

Step 3: Initialize learned weights of trained models

Step 4: Multiply weights by input and sum them up

Step 5: Compare result against the threshold to compute the output (1 or 0)

Step 6: Update the weights

Step 7: Repeat

Step 8: Predict Academic Performance with selected model as Output

Step 9: End

3.2.5 Methodology of the Study

Many system methodologies are available to structure, plan and control the overall activities involved in the optimization of the existing system or development of the new system. These methodologies combine set of principles, practices and processes that allows the development of systems quickly and properly. This research work on an Improved Hybrid model for predicting academic performance will be achieved following the Object-Oriented System Development Methodology (OOSDM). This is aimed at viewing, modeling and implementing the proposed system as a collection of interacting classes and objects. OOSDM is adopted because it is more effective, efficient, reliable, reusable and a faster way of developing systems. Furthermore, the Object-Oriented System Development Methodology (OOSDM) is a technical approach for analyzing and designing an application, system, or business by applying object-oriented programming, as well as using visual modeling throughout the software development process to guide stakeholder communication and product quality.

V. RESULTS AND DISCUSSION

TABLE 5.1: PERFORMANCE EVALUATION OF THE PROPOSED SYSTEM

(**Source:** SPSS processed DFGOS respondents input. SD = Standard Deviation)

S/n	Questions (Q)	Male end-user response		Female end-user response		Average Mean(x)	DFGOS automated remark
		Mean	SD	Mean	SD		
1	**Q1:** The DFG Model for predicting students' academic performance has a good Graphical User Interface (GUI)	2.87	1.29	2.44	1.51	2.75	Accepted
2	**Q2:** The DFG Model for predicting students' academic performance is fast in the validation of user access	2.75	1.23	2.31	1.43	2.28	Accepted
3	**Q3:** The DFG Model for predicting students' academic performance is accurate in the prediction process	2.73	1.92	2.63	1.65	2.29	Accepted
4	**Q4:** The DFG Model for predicting students' academic performance is fast in the prediction process	2.74	1.34	2.36	1.24	2.87	Accepted

5	**Q5:** The DFG Model for predicting students' academic performance is secured and flexible	2.65	2.99	2.78	1.12	2.57	Accepted
Grand Mean		**2.74**	**1.78**	**2.71**	**1.67**	**2.57**	

SPSS = Statistical package for social sciences, DFGOS = Deep Fuzzy Genetic Online Survey

TABLE 5.2: COMPARATIVE ANALYSIS OF THE EXISTING AND PROPOSED SYSTEMS RESULTS

Parameter	**Existing system**	**Proposed system**
	Grand Mean of user satisfaction	
GUI	2.59	2.74
Accuracy of the prediction process	0.78	1.78
Speed of the prediction process	2.68	2.71
Speed in user validation	0.67	1.67
Security	2.52	2.57

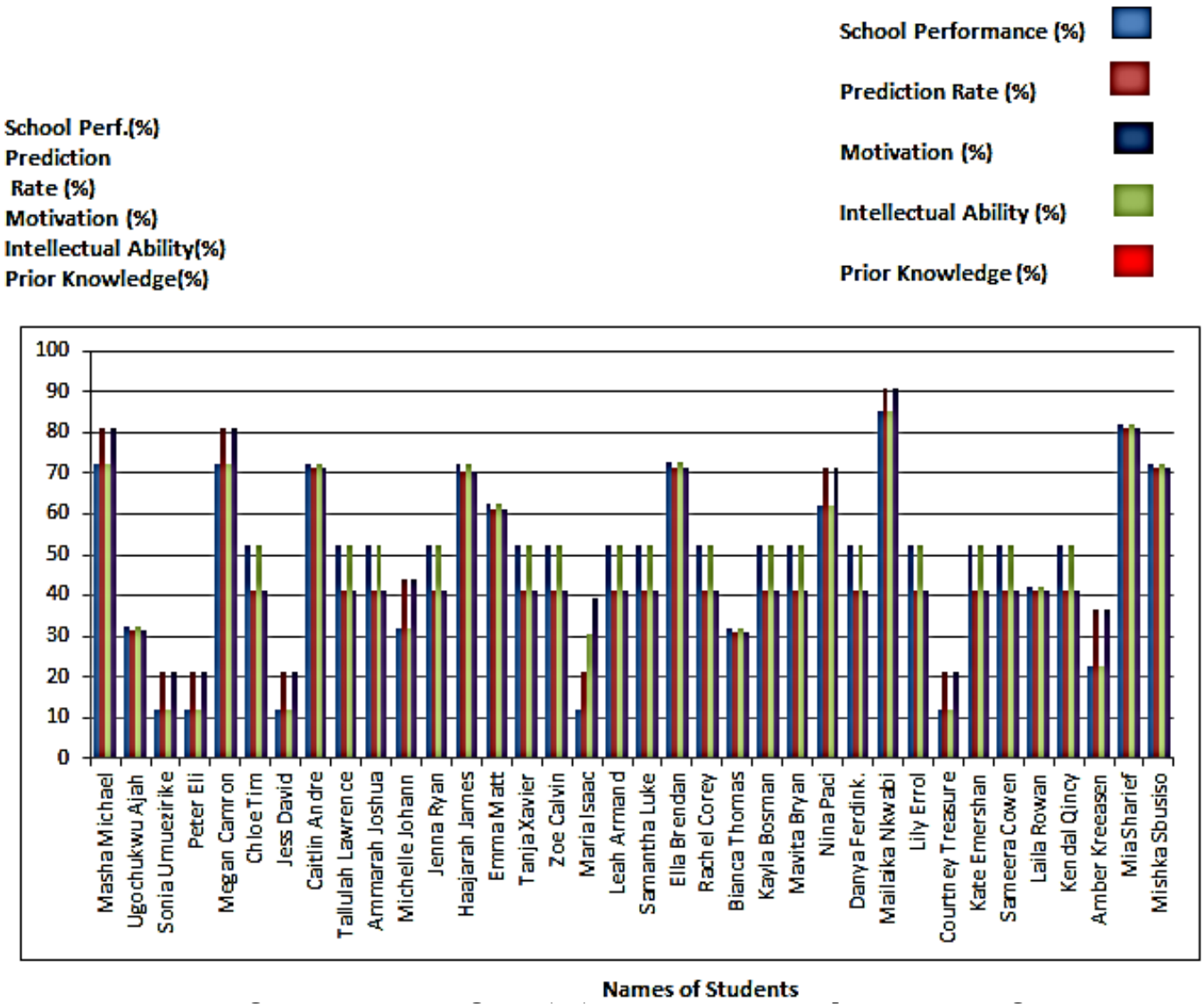

Figure 5.1: Chart representation of the proposed system result

5.1 Discussion of Results

Table 5.1 shows result of the newly developed Deep Fuzzy Neuro-Genetic model for the prediction of students' academic performance before admission into Tertiary Institutions in South-Africa and Nigeria (i.e integration of the existing and proposed system test-sets). The key indicators the newly developed model used in the prediction are school performance rate, motivation, intellectual ability, prior knowledge, prediction rate of academic performance, and the model remarks/interpretation. The system predicted students from 60% and above to be competent in academic performance, while students within the range of 40% to 50% were fairly competent in academic performance. Furthermore, the newly developed model utilized machine-learning oriented algorithms such as deep learning, fuzzy logic and genetic algorithm for training

and testing the model for accurate prediction of students' performance. Deep learning is a subset of machine learning in artificial intelligence (AI) that has networks capable of learning

Unsupervised data that is unstructured or unlabeled. Deep learning has evolved hand-in-hand with the digital era, which has brought about an explosion of data in all forms and from every region of the world. Fuzzy logic is a branch of machine learning that is being utilized to determine partial truth of a computing process using Boolean value which includes 0 and 1. Fuzzy logic algorithm helps to solve a problem after considering all available data. Genetic Algorithm is a machine learning algorithm that imitates the process of natural selection. It is used for solving complex search optimization issues.

Genetic Algorithm is a machine learning algorithm that imitates the process of natural selection. It is used for solving complex search optimization issues. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction. In addition, a Deep Fuzzy Genetic Online Survey (DFGOS) of the proposed system was designed and deployed on the World Wide Web (WWW) alongside the program in order for end-users to evaluate and give feedback on the system performance. The number of respondents to the designed DFGOS comprised of a Hundred and Seventy Three (173) of relevant stakeholders in Tertiary Institutions. Furthermore, the designed DFGOS contained two sections A and B. Section A elicited information on a demographic data of the respondents with three items namely the name of the school, designation and years of experience of the relevant Stakeholders, while section B elicited information from the prescribed survey questions. The section of the designed DFGOS is on a four point likert type scale as follows:

i) Strongly Agree (SA) – 4 points

ii) Agree (A) _ 3 points

iii) Disagree (D) – 2 points

iv) Strongly Disagree (SD) – 1 point

Also, the designed DFGOS was examined by experts in assessing students' academic performance. The construction of the DFGOS was also guided by relevant supervisors and experts in Tertiary Institutions, and the comment from the supervisors and experts in assessing students' academic performance were used to modify items in the questionnaire. To determine the reliability of the designed DFGOS, a robust web-based backend was interfaced with the frontend of the survey to enable rapid processing of the respondents' input on the prediction of students' academic performance before admission into Tertiary Institutions.

VI. CONCLUSION AND FUTURE WORK

In this study, we have developed an improved deep fuzzy neuro-genetic model for predicting the academic performance of students seeking admission. A major criterion for determining the eligibility of candidates for admission into Nigerian universities was their performance in UTME, WASSCE or its equivalent and Post-UTME. The study also aimed at investigating the performance of students in UTME conducted by Joint Admissions and Matriculation Board (JAMB) and the Post-UTME conducted by tertiary institutions in Nigeria. Furthermore, the reality of the present system failure is fully demonstrated by the students' performance in their final year. Also, there have been concerns about the quality of graduates being produced by

Nigerian universities. The study contributed the following improvements to predicting academic performance of admission seekers to tertiary institution in Nigeria:

i) An improved machine learning-based hybrid model for analyzing and integrating past records of students seeking admission into tertiary institutions.

ii) A decision-support model for the management of tertiary institutions especially in the selection of students for admission.

iii) A structured correlated framework for accurately predicting the academic performances in quick response time.

We intend to further expand the scope of the study in future to encompass the prediction of undergraduates that will be mobilized for national service. The proposed suggestion will enable the National Youth Service Corps (NYSC) adequately integrate data and further prepare for catering to large batches of potential corpers for youth service.

REFERENCES

Acharjya O. and M. Ming (2016), Data Management Concepts for Large Spatial Databases,

International Journal of Engineering Technology (IJET), 4(5), 1 - 4

Ahn .B. and M. Nelson (2017), Info Data for the Telecommunications Industry, *An*

International Journal of Computer Applications (IJCA), 3(2), 24 – 32

Badra, G., Algobaila, A., Almutairia, H and Almuterya, M. (2016). Predicting Students'

Performance in University Courses: A Case Study and Tool in KSU Mathematics Department. *Symposium on Data Mining Applications, SDMA2016.82*, pp. 80-89. *Riyadh, Saudi Arabia: Elsevier.*

Bamnote G.R., S.S Agrawal (2013), Introduction to Query Processing and Optimization,

International Journal of Advanced Research in Computer Science and Software Engineering, 3(7), 53 - 56

Borde S., Agrawal U.M., Bilay V.S. and Dogra N.M. (2017) Supervised Machine Learning techniques for Education Data Mining, *IJSART, 3 (3), 760-764*

Chandrashekhar W. G, Gaurav V. W, Mohit M. D. and Ganesh V. P. (2015) Clustering Based Feature Selection Algorithm. *IJCSN International Journal of Computer Science and Network, 4, (2), 346-350.*

Chen, H. (2016). Applications of Fuzzy Logic in Data Mining Process. In Z. H. Bai Y., *Advanced Fuzzy Logic Technologies in Industrial Applications. Advances in Industrial Control.* Springer: London.

David R.O., Florentino F.R. and Jose R.M. (2018) Using Evolutionary Computation for

Discovering Academic Data Prediction, *Elsevier Journal, Information Processing and Management 54, 303-317.*

Dhinakaran V. (2019), StealthDB; a Scalable Encrypted Database with full SQL query Support,

Proceedings on Privacy Enhancing Technologies; 2019(3), 370 – 388

Dixin T., S.Zechao, E.Aaron (2016), Intermittent Query Processing, *Journal of Computer and Software Research (JCSR), 10(8), 144 - 147*

Dominik D. (2019), Experimental Study of Memory Allocation for High Performance Query Processing, *International Journal of Computer Application (IJCA), 14(7), 17 - 22*

Eke N, G. Twen (2017), Info. Data concepts in Data Management, *International Journal of Computer Applications (IJCA) 3(2) 20 - 25*

Federal Republic of Nigeria (2006), National Policy on Education, *NERDC, Abuja, Nigeria, 1 - 7*

Gbolagade, M.D., Hambali, M.A and Akinyemi, A.A. (2015), Predicting Postgraduate Performance Using Resample Preprocess Algorithm and Artificial Neural Network. *African Journal of Computing & ICT, 8(1), 145-158.*

George P.C. and Savaliya B.R. (2017) Sender Recognition of Search Data Categorization, *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 6 (6), 107-111*

Ghosh A. and Pasayat A.K. (2018) Identifying Data Prediction using Apache Spark MLlib, *Journal of Emerging Technologies and Innovative Research (JETIR), 5 (5), 893-897*

Giatrakos .G. and N. Hillong (2018), Database Systems: *An International Journal and Article for the Department Of Computer Science, University of Stanford, 1 - 6*

Hamsa, H., Indiradevi, S and Kizhakkethottam, JJ. (2016). Students Academic Performance Prediction Model using Decision tree and Fuzzy Genetic Algorithm. *Global Colloquium in Recent Advancement and Effectual Research in Rngineering, Science and Technology* (pp. 326-332). Elsevier.

Henry W., M. Vukosi and M. Sengeh (2019), Predicting School Performance Using a Combination of Traditional and Non-Traditional Education Data, *International Journal of Science Research (IJSR), 14(9), 1 - 6*

Ibrahim (2018) Hybrid Approach to Predictive Data using Preventive and Curing Techniques, *Journal of AL-Qadisiyah for computer science and mathematics, 10 (3), 16-24*

Ihab I. (2016), Supporting Top-K Join Queries in Relational Databases, *International Journal of Engineering Technology (IJET), 4(7), 131 - 137*

Ijaiya N. (2018), An Investigation into the problems of teachers' management in Kwara State Secondary Schools, *Studies in Educational Planning and Administration, 1(2), 49 – 58*

Imran, M., Latif, S., Mehmood, D and Shah, M.S. (2019), Student Academic Performance Prediction using Supervised Learning Techniques, *International Journal of Emerging Technologies in Learning (iJET), 14(14), 92-104*

Jabbarifar, T. (2009). The Importance of Classroom Assessment and Evaluation in Educational System. *2nd International Conference of Teaching and Learning (ICTL 2009)*, (pp. 1-9). Malaysia.

Lassa P. (2010), Teacher Production: A Focus on Nigeria, the State of Education in Nigeria, *UNESCO Abuja, 70 – 83*

Laxman J.K. and Khansole B.A. (2017) Machine Learning Approach for Academic Data Prediction, *International Journal of Innovative Research in Computer and Communication Engineering, 5(7), 13114-13121*

Lukong, E.T., Njamfa, T.N and Michael, M.E.E. (2017). Redefining Educatioin in Africa, "The need for Sustainable Development through Innovation and Creative Learning". *International Journal of Trend in Scientific Research and Development (IJTSRD), 1*(6), 217-236.

Macdonald P. and K. Lings (2015), Types of SQL Database and its comparison with Info, Data

Management, *International Journal of Applied Information Systems (IJAIS) 5(4), 12 – 19*

Oluwaseun O., Emmanuel O., and Olutosin O. (2017), Analysis of Candidates' Performance in Unified Tertiary Matriculation Examination (UTME) and Post-UTME in the University of Ibadan, Nigeria, *Journal of Scientific Research and Reports, 14(5), 1 – 8*

Omisore, M.O; Samuel, O.W and Atajeromavwo, E.J. (2017). A Genetic-Neuro-Fuzzy inferential model for for the Diagnosis of Tuberculosis. *Applied Computing and Informatics, 13*, 27-37.

Osegi, N.E. (2015). *A Generative Model for Multi-Dialect Representation, 1-16*

Owolabi S.O (2017) Teacher Education in Sub-Saharan Africa: Policies and Practices, *Paper Presented at the Development of Teacher Education in Africa Conference, Makarere University, Kampala*